DOCUMENT RESUME

ED 082 861 PS 006 965

AUTHOR Webb, Roger A.
TITLE Studies of Young Children's Thought and Language. Final Report.
INSTITUTION Johns Hopkins Univ., Baltimore, Md.
SPONS AGENCY National Center for Educational Research and Development (DHEW/OE), Washington, D.C. Regional Research Program.; National Inst. of Mental Health (DHEW), Rockville, Md.
BUREAU NO BR-2-C-031
PUB DATE Jul 73
GRANT OEG-3-72-0048
NOTE 71p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Cognitive Development; *Concept Formation; *Language Development; *Linguistics; *Preschool Children; Psycholinguistics; *Semantics; Thought Processes

ABSTRACT

A series of studies on young children's use of the terms "same" and "different" are reported. The work began from the observation that young children could respond correctly to instructions involving "same" but were often incorrect in response to "different". This finding was replicated under a variety of experimental conditions and found to be reliable and stable. Same-different judgement tasks were found to be unsuitable for use with young subjects, and no significant correlation between this same-different phenomenon and classification performance was found. The findings are reviewed in terms of several linguistic models, and a four stage developmental model for the child's use of different" is postulated. The significance of these findings for a theory of semantic development and for experimental and instructional work with young children is discussed. (DP)

Final Report

Project No. 2-C-031
Grant No. OEG-3-72-0048

Roger A. Webb
The Johns Hopkins University
34th and Charles Streets
Baltimore, Maryland 21218

STUDIES OF YOUNG CHILDREN'S THOUGHT AND LANGUAGE

July 1973

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of Education

National Center for Educational Research and Development

(Regional Research Program)

Abstract

A series of studies on young children's use of the terms "same" and "different" are reported. The work began from the observation that young children (approximately three-and-a-half years of age) could respond correctly to instructions involving "same" but were often incorrect in response to "different." This finding was replicated under a variety of experimental conditions and found to be reliable and stable. Same-different judgment tasks were found to be unsuitable for use with young subjects. There was no significant correlation between this same-different phenomenon and classification performance. After considering several linguistic models of the phenomenon, a four stage developmental model for the child's use of "different" was postulated. It was postulated that the youngest children (approximately 3-2 and younger) reversed the meaning of "different" for "same." Children approximately 3-3 to 3-7 believed that "different" meant a different example of a similar class of objects (i.e., a denial of identity). Older children 3-7+ apparently believed that "different" meant different with some dimension of similarity. Still older children finally appeared to arrive at a formulation comparable to that held by adults. The significance of these findings for a theory of semantic development and for experimental and instructional work with young children is discussed.

Final Report

Project No. 2-C-031
Grant No. OEG-3-72-0048

STUDIES OF YOUNG CHILDREN'S THOUGHT AND LANGUAGE

Roger A. Webb

The Johns Hopkins University
34th and Charles Streets
Baltimore, Maryland 21218

July, 1973

The research reported herein was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
National Center for Educational Research
and Development

Preface

The work reported here was supported in its earliest phases by a grant from the National Institute of Mental Health (MH-19364). The author wishes to thank Mary Ellen Oliveri, Francis Harnick, Susan Rose, Linda O'Keefe, and Sally Zulver for their invaluable assistance with various phases of the research. Subjects were obtained with the co-operation of the Grace and St. Peter's Episcopal Church School, Brown Memorial Presbyterian Church School, the Cathedral of the Incarnation School, and the Baltimore Ethical Society School. The assistant of these organizations is gratefully acknowledged.

Table of Contents

Table of Contents, Continued

List of Tables

Introduction

The following report covers a series of studies of young children's use and understanding of the terms "same" and "different." The work grew out of three considerations. 1) The investigator worked to find an area of content that would allow him to study the interaction between language, cognition and social input. Children's notions of similarity and discrepancy seemed to be ideal since there was a literature on the growth of operation of classifications (e.g., Inhelder & Piaget, 1964; Kofsky, 1966) as well as literature on the development of relational terms (e.g., H. Clark, 1970; Donaldson & Balfour, 1968; Griffiths, Shantz, & Sigel, 1967). 2) The immediate impetus for the following studies was the discovery of Donaldson & Wales (1970) that young children appeared to be almost completely wrong in their responses to instructions involving "different," suggesting that they revised the meaning of the term in some sense. 3) Finally, the terms "same" and "different" were viewed as particularly significant because of the relationships they marked. There is a deep and abiding controversy in psychology on the nature of psychological <u>similarity</u>. Psychometricians (e.g., Torgerson, 1965) have been concerned with scaling stimulus-based notions of similarity. The present writer, however, is more in agreement with another tradition which believes that similarity is inevitably <u>post hoc</u> (e.g., Elkind, 1969,

p. 173 ff.). The latter point of view is that judgment of similarity involves classification and that the potential bases of classification are infinite. Thus, for any two objects, an intelligent human can reasonably justify either the conclusion that they are the same or that they are different and it will be impossible to disagree. People do not live in a constant quandary about such questions, though, because there is general consensus within a cultural group over what one implies when one asks a question about similarity--which is also, of course, what makes the scaling of similarity possible. The key aspect of this problem for the developmentalist is that the referents for the terms "same" and "different" are inherently social. People can only talk about these matters to the degree that they share common systems of categories. Thus judgments about similarity are ideal subject matter for studying the role of social influence on intellectual development.

It should be noted that the most popular perspectives on cognitive and linguistic development minimize the role of social factors. Piaget's classic studies suggest that each child must reinvent the cognitive wheel for himself. Where social influences do come in (e.g., the rules of a game, Piaget, 1965) the child accommodates to the social reality rather than being shaped by it. In a similar fashion most recent work on language views the developmental process as an unfolding along predetermined lines in which the child maps his own innate sense of language

against the language being spoken in his community. Of major writers in the field, perhaps only Vygotsky (1960) and more recently Bruner (Greenfield & Bruner, 1969; Cole & Bruner, 1971) have consistently advocated a central role for social influences in intellectual development.

The report which follows is in the form of three studies. Each has its own independent introduction, method, results, and discussion sections. There is then a final discussion section in which overall conclusions are reported. Two points must be made before turning to the studies proper. First, the sequential development of the work reflects a changing attitude on the part of the investigator as to the true nature of the phenomena under study. Initially, the investigator believed that, although it was terribly interesting, the basic reversal phenomenon discovered by Donaldson & Wales [hereafter D & W] was an artifact of the experimental procedure and would not be supported by further investigations. Specifically the investigator anticipated that manipulations of the choice set would eliminate the phenomenon altogether, and more productive questions could then be studied. Just the opposite has been found. The phenomenon discovered by D & W has proved to be most reliable and stable across a number of variations in the basic procedure.

A second point which must be raised concerns the use--or non-use--of statistics. The initial study to be reported below underlines the problems of conventional statistics in the analysis of developmental

data of the sort reported here. Often the most interesting questions concern the means employed by individual subjects in solving the problems presented to them. If an overall analysis such as an analysis of variance is employed, variations in individual response strategies become lost in the error term. It is impossible to tell from such an analysis whether individual variation is error, a different approach to the problem, or perhaps a developmentally distinct phase in the child's intellectual growth. For these reasons, the stress in the results sections which follow is on the description of individual patterns of behavior. Fortunately, the patterns of responses seen were far from infinite and in fact were quite regular. It was, thus, possible to describe the data in relatively few basic patterns. There is, of course, no guarantee that the results from the samples employed in the present studies will generalize to other samples, but statistics will not increase that assurance. On the samples studied, the basic findings are very stable, replicable, and, in one longitudinal study, consistent over time. Unless the samples under scrutiny are unique, one could reasonably expect that the results to be reported below would be easily replicated with other samples.

The first of the studies to follow deals with an attempt to replicate D & W's reversal phenomenon with a non-choice technique. This experiment failed, but for very interesting reasons and supports the methodological decisions made in the remaining studies and discussed above. The second study established certain basic facts about behavior

in the D & W procedure and the systematic nature of the behavior observed. The third study is an attempt to establish a sequence of stages in the development of the meaning of "different" and to examine further the relation of the child's language to non-linguistic choice behavior in the D & W task.

Study I:

How children do not answer questions: The occurrence of response-generating heuristics in young children's judgments of similarity and difference

Introduction

The present study grew from an attempt to study young children's understanding of the term "different." Using a task that required a child to select an object that was either "the same in some way" or "different in some way" from a target, D & W found that children of about three-and-a-half years were correct in response to the "same" request, but also chose similar objects in response to "different." D & W interpreted these results as indicating that "different," in this age group, might refer to a nonidentical object with similar qualities. This could occur since in adult English the terms "same" and "different" refer to relations both of identity and similarity. Because of the methodological thrust of the following report, the theoretical implications of D & W's findings will not be discussed at length. The present writers assume, however, that: 1) the same-different reversal found by D & W is interesting for a theory of semantic development, and 2) D & W's proposed explanation is one of several plausible explanations for the phenomenon.

The study to be reported began as an attempt to replicate D & W's findings with a technique radically different from the choice task they had employed. It was assumed that the use of a judgment task in

which the objects to be judged were simultaneously present would eliminate many of the possible interpretations of the D & W effect, including explanations based upon the confusion of similarity and identity or possible explanations based upon information processing models.

The research plan consisted of a manipulation of question form and of the type of material to be judged. Upon completion of this phase of study, it was found that the question form and material variables plus a classification of subjects by age produced highly significant effects. A closer inspection of the data, however, indicated that the children were not dealing with the questions in the expected fashion. Briefly, subjects appeared either to answer the judgment questions correctly or to use an heuristic that allowed them to evade the problem altogether.

The analysis required to discover the true nature of the results was unusual; for this reason, the authors wish to present this study as an example of a subject-by-subject analysis of group data. A feature of the analysis is the use of a stringent statistical criterion for attributing strategy-use to subjects.

METHOD

Subjects. A total of 106 children ranging in age from three years, one month to five years, eleven months were studied. Children were enrolled in four nursery schools in the Baltimore area and the sample was predominantly middle-class. Approximately equal numbers of boys and girls were studied. All the children of appropriate ages that were available in the schools were tested.

Materials. Two sets of test material were employed. Experiment 1 involved

the children's judgments of pictures. Two parallel sets of 16 pairs of pictures or paper constructions were mounted on 10 X 12 inch cards. Four of the pairs consisted of pictures of common objects (e.g., a wagon or a lawnmower), four consisted of pictures of people, four were geometric shapes of varying colors (squares and triangles, red and blue), and four were patterns of three small colored squares set on a larger square representing block designs. Of each four, two of the pairs were the same and two were different. Of the two different geometric shapes, one pair differed on both dimensions and one pair differed on only one dimension.

Experiment 2 involved the judgments of three-dimensional objects. Again, two parallel series of 16 pairs were employed with half being identical and half differing on at least one visible dimension. Eight pairs of geometric shapes were made up of triangles, squares, and circles, which were red, green, blue, or white. Of the four pairs of geometric objects which differed, two differed in both color and shape and two differed on only one relevant dimension. The remaining eight pairs consisted of common objects (e.g., toothbrush, small dolls, etc.).

Procedure. A common procedure was used for both Experiments 1 and 2. To familiarize the subject with the material, a sample card with two geometric constructions was employed in Experiment 1, and in Experiment 2, two geometric objects. The experimenter asked the child about the names and colors of the objects. After familiarization, the experimenter asked the child about the sameness or differentness of the sample pair using the question form which was to be used in the remainder of the procedure.

Three question forms were employed: (1) "Are these the same?" (2) "Are these different?" (3) "Are these the same or different?". These three forms in combination with the two parallel sets of pairs generated six experimental cells to which the children were randomly assigned.

A second experimenter recorded the child's verbal response and any additional remarks that were made. The entire session was also tape recorded. Often a third adult was present who was either a second observer, a teacher, or other interested adult. After the child's answer to each question, the experimenter nodded his head and said "okay, "good," or "fine," with an effort at equal enthusiasm regardless of the child's answer.

RESULTS

Almost without exception, the grammatical form of the children's answers was appropriate to the questions. That is, the questions, "Are these the same?" or "Are these different?" were answered "yes" or "no," while, "Are these the same or different?" was answered "same" or "different." Only two children in the "same or different" group answered questions with "yes-no" answers, and these children were dropped from the ensuing analyses. A few "same" and "different" answers were scattered in apparent random fashion throughout the conditions in which "yes-no" answers would have been more appropriate. These were scored and used, however, since their interpretation was unambiguous.

Several preliminary analyses of the data were conducted using multivariate analysis of variance (MANOVA). There were no significant effects

of the two parallel series and data were then collapsed across series. Experiments 1 and 2 were found to be significantly different ($p < .001$) but examination localized the source of difference in one type of item (c.f. below). To avoid confusion, data from the two procedures were analyzed separately. Finally, analyses using different possible dependent variables indicated that the best analysis resulted from using the number of correct responses on the eight same and eight different pairs as two dependent variables. For further analysis, data recorded during the sessions were transcribed as being either correct or incorrect, thereby eliminating the differences in actual responses (i.e., "yes," "no," "same," or "different").

For Experiment 1, the overall MANOVA indicated that there were significant effects of age ($p < .001$), question ($p < .005$), and age by question interaction ($p < .013$). The comparable analysis of Experiment 2 in which objects were employed indicated that there was a highly significant effect of question form ($p < .001$). The age effect, however, only approached significance ($p < .086$), and the age by question interaction was not significant ($p < .746$). Overall, the judgments of same and different pairs showed different patterns with a low negative correlation between the two variables ($r = -.20$).

The major question suggested by the D & W result (c.f. above) was whether there were any subjects who were completely wrong in the "different" question condition. In examining the data with regard to this question, four such subjects were found, but many other subjects were discovered to be dealing with the experimental materials in ways other than that intended by the experimenter. Therefore, an attempt was made to analyze exhaustively all 106

subjects for the heuristic they were using in the task. Because of the difficulties of analyzing subjects who did not make a number of errors, the decision was made to assume that all children with a number of correct responses above chance (the probability of 12 or more out of 16 is .038 assuming a binomial distribution with p = .5) were responding to the task correctly. A total of 68 subjects met this criterion and were not available for further analysis. Also, as was mentioned above, two subjects were dropped because of ambiguous answers. There were thus 36 subjects available for an analysis of the heuristics used to generate answers.

Four of the 36 subjects were significantly below chance with two or three correct answers, most of which were on items that other subjects missed. All of these children were found in the "different" question condition. These four were the only children of 32 in the "different" condition who responded in a fashion reminiscent of the D & W results and apparently reversed the meaning of "different" for "same." These children were not, however, all near three-and-a-half years, but were scattered across the age range.

The remaining 32 children fell in the chance range (5-11 total correct responses) but were not necessarily responding on a chance basis. The deviation of the outcome from chance was determined by considering each subject's answers in a 2 X 2 table in which same and different pairs were on one dimension, and correct and incorrect on the other. Thus, a subject with a "chance" outcome of eight correct responses may have all eight on the same pairs and none on the different pairs. Each subject's outcome was tested using Fischer's exact test (Hays, 1963).

Fourteen of the 32 subjects obtained significantly more of their correct responses on different pairs, and eight of the 14 had a perfect 0-8 split (i.e., zero correct on same, eight correct on different). In ten cases, there were significantly more correct responses on same pairs, with three subjects showing a perfect 8-0 split. The remaining eight subjects were apparently at chance levels with correct response outcomes of 5-5, 3-3, 2-4, 2-6, 5-1, and 1-4 on same and different items. None of these outcomes can be called non-random with a probability of less than .05.

To complete the analysis of the heuristic employed by the subjects, it was necessary to translate the correct and incorrect answers back into the actual responses. This transformation indicated that of the 15 subjects in the "same" and "different" question conditions who used an identifiable heuristic, 14 were saying "yes" to all or nearly all items. One child in the "different" condition answered "no" to all items. Of the nine subjects who got most of their correct answers on one type of pair in the "same-different" condition, six answered "different" most of the time, and three, "same." Summarizing the individual analyses, it can be seen that of 36 subjects at or below chance levels of responding, four may have reversed the meaning of "different" for "same," 14 simply answered "yes" to most questions, and one "no," six usually answered "different," and three, "same." Eight subjects appeared to be guessing.

After the analysis above confirmed that it was possible to describe the strategies used by subjects at or below chance levels, the MANOVA's were repeated with these subjects removed from the sample. The deletions left 34 subjects in each of Experiments 1 and 2. Since younger subjects predominated in the groups dropped from the analysis, it was impossible to use age as a factor in the design. Therefore, age in months was used as a

covariate and a one-way MANOVA was performed on the question effect. The two experimental procedures were analyzed separately, as before, and the total correct on same pairs and the total correct on different pairs were again used as the two dependent variables. The analysis performed in this fashion revealed that there were no significant effects left in the data after the removal of the deviant subjects. Neither the within cells regression with age nor the question effect was significant in either Experiment 1 or Experiment 2. The regression effect of age in Experiment 1 approached significance ($F=2.46$, $p < .10$) but all other F's were less than 1.0. When the comparable analysis of covariance was performed on the full set of data, the within cells regressions were significant and question effects were highly significant in both Experiments 1 and 2. Thus, with the full set of data, the analysis of variance using age as a factor and the analysis of covariance using age as a covariate gave equivalent pictures of the results.

An item analysis of the materials used in Experiments 1 and 2 proved to be unremarkable with one exception. The difference in overall difficulty between Experiments 1 and 2 (see above) appeared to be due almost entirely to the different block design pairs in Experiment 1. These were called "same" by most subjects when those following one of the heuristics described above were excluded. The proportion of correct responses on these items was .07 compared with an overall proportion of .92 with these items excluded. With this exception, there would have been no significant difference in overall difficulty between the two procedures. Some items were more

difficult than others even for the children who were above the chance criteria, but these will not be discussed in light of the overall findings.

DISCUSSION

The direct contribution of the present study to our understanding of the young child's notion of "different" would appear to be marginal, though the reasons for this negative conclusion are of some interest. Although the grammatical form of responses to the same-different judgment questions was correct, the answers given by children who made a considerable number of errors could not actually be considered answers to the questions asked. The sample could easily be divided into a group of 68 which performed the task more or less correctly and a second group of 36 which did something altogether different from the experimenter's intentions. Within the group of 68 subjects, there were no significant effects of question form or age. The remaining 36 subjects who gave interpretable responses could be characterized as using one of a number of possible heuristics for dealing with the problem--the most common being to answer "yes" to all questions. Several subjects were presumed to be guessing after their performances were found actually to be at chance levels.

One virtue of the present analysis would appear to be the strong evidence for the heuristics which are postulated. For each subject considered individually, the probability is less than .05 that he deviates by chance alone from random responding. The only exceptions to this were the eight subjects presumed to be guessing. While these patterns might be called random by some, the present writers prefer to consider guessing a legitimate heuristic which is adaptive in situations of high uncertainty.

The experimental variables produced highly significant results when the full sample was analyzed, but no significant effect when the subjects using deviant strategies were removed. The drop in significance would not appear to be due to the loss of degrees of freedom since with one exception the obtained F's fell below 1.0. Thus, the experimental variables would appear to function by changing the frequencies of subjects employing a non-standard heuristic. It is possible, of course, to discuss the relative difficulty of the various questions used, but the appropriate analysis would require frequency or proportion statistics rather than statistics based on the means of the variables.

The incidence of heuristics which, in effect, allow the child to avoid the problem posed by the experimenter invites one final speculation. Accounting for the observed results would seem to require the recognition of the child's acceptance of a role. By agreeing to "play" with the experimenter, the child accepts the responsibility to respond to questions. When the questions are uninterpretable or unanswerable, the child must find some device for fulfilling his obligations as a subject. In the present case the child generates answers which are grammatically correct, but not, strictly speaking, answers to the questions posed. The high proportion (roughly one-third) of subjects responding in this fashion on such a simple task suggests that similar behavioral strategies may confound the interpretation of other studies. If analyses end with an analysis of variance, such interesting findings (or serious problems, depending on the point of view involved) might go unnoticed. To the present writers, finding out what question the child is answering may be the heart of developmental research.

Study II:

Thought and language interaction in children: Investigations of the Donaldson and Wales Same-different phenomenon

Introduction

One of the more striking phenomena in young children's linguistic development is their occasional confusion of a word for its opposite. Vygotsky (1960, p. 71 ff.) noted this in his classic Thought and Language, and Donaldson and Balfour (1968) have brought the fact to our more recent attention. Little mistakes often provide exciting clues to the nature of intellectual growth--as Piaget's many studies demonstrate--and the child's reversal of paired terms may be just such a happy error. It is not surprising, therefore, that several recent studies (e.g., E. Clark, 1971; Donaldson and Balfour, 1968; D & W, 1970; Gordon, 1972) have focused on children's confusions of antonymic adjective pairs.

The theory of lexical marking (e.g., Greenberg, 1966) assumes that the meaning of a term is composed of a set of features which are added in quantum fashion to the mental dictionary entry for a lexical item. In this theoretical context, H. Clark (1970) argued that as the meaning of, for example, "big" and "little" develop, both refer nominally to physical extensivity and only later are differentially marked for comparative use. Thus, the young child might confuse "little" for "big" when both imply only that there is some bigness to the things being discussed. "Big" is the unmarked member of the adjective pair

and "little," the marked, since the latter has an extra feature and is semantically more complex. E. Clark (1971) has further postulated that features are added in a fixed hierarchical order from superordinate to subordinate as children learn the meaning of terms.

D & W reported data which indicated that "same" and "different" followed a developmental pattern similar to that seen for antonymic adjective pairs. Using a task which required three-and-a-half-year-old children to select objects that were either "the same in some way" or "different in some way" from some target object, D & W found that children responded correctly to the "same" request, but also selected a similar object in response to "different." This occurred when the objects in the choice array were either just like or completely different from the target object. In a second condition of the D & W experiment, the objects in the choice array all shared one dimension with the target and differed on another, and the child could not make a clearly erroneous choice to either request. D & W's explanation of the phenomenon was based on the inherent ambiguity of the English language. The terms "same" and "different" can refer either to similarity relations or to identity relations, and the exact meaning must be inferred from either the physical or linguistic context. "Different," D & W suggested, along with other possibilities, might initially refer to a different example of a similar object. When the child selects a similar object he is saying, in effect, here is another member of that class.

Clark (1970) suggested that D & W's hypothesis could be tested by returning the target object to the choice array and seeing whether the child chose the identical target of another, similar, one. If D & W were correct,

the child should never choose the identical object when asked for one that is "different." Unfortunately, there does not seem to be an unambiguous interpretation of Clark's proposed experiment even if the technique were to generate clear data. It is the present writers' intuitions that replacing the object changes the context of the remarks and implies that the non-identity choice is anticipated by the experimenter. Also, replacement changes the task by introducing a memory factor which is not present in the original D & W experiment, i.e., the child must remember both the properties and the location of the returned token. Thus, the task does not appear comparable to the original D & W experiment.

Regardless of the difficulties in clarifying D & W's findings, the data are extremely interesting for a theory of semantic development. The reversal bears at least superficial similarity to the data from antonymic adjective pairs although "same" and "different" deviate from such adjective pairs in several ways. "Same" and "different," in a strict sense, are not adjectives but are the linguistic markers of class membership and object permanence. They refer ambiguously to relations of both similarity and identity and there is no physical dimension which is their obvious referent.

Caution must also be observed in making the interpretations of the D & W results. Both D & W and Clark assume that the child's deviant behavior follows directly from the structure of his semantic space. Many other factors could explain the findings, and one is not justified in concluding that some anomolous behavior in response to linguistic instructions follows directly from a misinterpretation of the words without carefully eliminating alternative explanations. In fact, the robustness of many developmental phenomena across a variety of linguistic instructions (e.g., the conservation tasks) suggests that

purely linguistic factors are unlikely to provide an adequate explanation of cognitive developmental phenomena.

In an attempt to examine some of the possible alternatives to semantic interpretations of D & W's results, in Study I the writer examined three- to five-year-old children in a series of same-different judgment tasks. Requiring the child to make judgments of simultaneously present objects with a number of linguistic instructions offered to eliminate several of the potentially correct interpretations of the results. Unfortunately, a subject-by-subject analysis of the judgment data indicated that the children were either almost completely correct on the task or used response-generating heuristics (e.g., saying "yes" to all questions) that allowed them to avoid the problem altogether. Thus, it appeared on the basis of the judgment task data that further attempts to understand D & W's results would have to be conducted with tasks more like that employed in the original D & W study.

The present investigations were undertaken to answer four basic questions: 1) Are D & W's findings replicable, and, if so, under what conditions? 2) Are the particular terms used critical or can a variety of terms produce similar results? 3) Is the effect systematic? Do children showing the D & W effect respond on the basis of a consistent rule which generates the choice of similar objects in response to the request for different ones? 4) Is there systematic individual variation in the rules used to generate choice behavior? Experiment I below was designed to examine the role of the specific words employed (question 2 above) and Experiment II attempted to investigate the internal consistency of the children's behavior (question 3). Both experiments, of course, are relevant to the replicability and generality of the findings

(question 1). Finally, the authors adopted a subject-by-subject analysis of the data in order to facilitate the detection of individual response patterns (question 4).

Experiment I

Subjects: Forty-nine children ranging in age from three years, one month to five years, seven months were studied. Taken from two Baltimore area private pre-schools, the sample was middle-class and predominantly white. All available children in the schools were studied.

Procedure: Each child was studied individually in a session which included tests of classification as well as the procedure to be reported here. After familiarization with the experimenters, and following half the classification testing, the child was introduced to a set of 12 objects: three pink plastic toothbrushes, three yellow plastic cups, three blue plastic combs, and three Sunshine raisin boxes. The child was asked the name and color of the objects and was allowed to play with them. He was encouraged to put the objects into classes and was prompted to call them by a full descriptive name (e.g., "That's a yellow cup."). In testing, the experimenter first removed one object from the set and asked, "Will you pick one for me that is different from this?" After the child's selection, the objects were returned to the set and the procedure was repeated four times with different questions. The child was asked for objects that were "same," "like," "not like," and again "different" in that order. During the session, the experimenter attempted to be supportive, but without giving differential reinforcement for any particular choice the child made.

Results: Of the 49 subjects tested in this fashion, 35 produced patterns of results that were easily classified. Nineteen were totally correct (i.e.,

always picked a different object when the question was "different" or "not like" and always picked a similar object when the question was "same" or "like"). Sixteen were perfectly consistent with the D & W pattern of results (i.e., they always picked a similar object regardless of the question). Five other children satisfied the criterion used by D & W in reporting their results and picked a similar object in response to the initial "different" request. If these five are added with the 16 who show consistent D & W patterns, there are a total of 21 D & W subjects in the sample of 49. Only one of these five subjects missed both "different" items while being correct on "not like."

If the 40 subjects who are either correct or satisfy the D & W criterion are considered, the results show an imperfect relationship with age. Splitting the sample at the median age of four years, four months generates a X^2 of .93 which is not significant. The best split that can be obtained comes from breaking the age dimension at four years and seven months. This leaves three D & W subjects above 4-7 and eight correct subjects below ($X^2 = 6.53$, $p < .025$). It should be noted, however, that two of the youngest subjects were completely correct and two D & W subjects were above five years of age.

The nine subjects who did not show one of the patterns described above require close scrutiny. Five of the nine made one or more errors that did not include a choice of a similar object in response to the initial "different" question. Four of the five made only one error and should probably be classified with the subjects who were correct on all items. The one remaining subject selected a different object as a response to every question.

The remaining four subjects were the most interesting observed. These children were correct when asked for a similar object, but when asked for one

that was "different," or "not like" the target, denied the existence of a satisfactory choice. Upon prompting with the questions, "What are you looking for?" or "What would you need to have a different one?" these children volunteered the opinion that they needed a similar, but not identical, object (e.g., "I need a pink comb," when the target was a blue comb).

Discussion

The outcome of Experiment I can be considered a replication and extension of D & W's results. Although the procedure deviates in several ways from the original D & W procedure, we found many children in the age range studied who would select an object similar on all dimensions to a standard when asked for something that was "different" or "not like"; also, these choices appeared to be independent of the particular terms employed. Only two subjects responded differentially to "different" and "not like."

It is not surprising that our results were somewhat less consistent than D & W's. Of our whole sample, about 40% (21 of 49) meet the D & W criterion while D & W found a higher percentage. Fifteen subjects generated 25 "same" responses out of a possible 30. Our sample, however, included a wider age range, and if we look just at children of less than four years (who would be more comparable to the D & W sample) the proportion of D & W subjects is .69. In addition to the age factor, the D & W subjects' uniform experience as participants in an experimental nursery school might contribute to the relative homogeneity of D & W's findings.

Unfortunately, the outcome of Experiment I did not give much information about the semantic space of "different." The subjects showing the D & W pattern could believe, as D & W suggested, that "different" initially refers to a denial of identity as opposed to similarity; they could confuse the

meaning of "different" for "same," (as D & W also considered); or they could be confused due to the way in which they process the information contained in the problem. The four subjects in the present sample who actually did volunteer information about the semantics of "different" suggested a third formulation. By saying that for two objects to be different they must also have something in common, the children suggested a theory of "different" that requires some observable basis of comparison. If a child were holding this theory, things that were different on all dimensions would not be different at all.

Experiment II

Introduction: Since Experiment I demonstrated the replicability of the D & W phenomenon, but did not satisfactorily differentiate between the various explanations for the finding, further study was designed. In D & W's original study a second set of objects had been used in which no identical object appeared, but shared dimensions of form and color were present. The present study attempted to combine the two D & W procedures while examining other relationships in a single procedure. In addition, we hoped to expand the children's explanations of their behavior by explicitly asking for justifications. A younger sample was also employed since more D & W subjects would be found.

Subjects: Twenty-nine children ranging in age from two years, nine months to three years, eleven months were studied in two church-operated preschools in Baltimore. The sample was predominantly white and middle-class. A few days to a week prior to the individual testing, the experimenters visited and played with the children in the group setting.

Procedure: The experimenter sat next to the child at a small table, with one or two observers present. The entire session was tape-recorded. Each child was presented with an array of eight objects: two red, flat plastic squares (approximately one square inch in size); one white plastic comb; one blue plastic ring (approximately 1-1/2 inches in diameter); one yellow unsharpened pencil; one green unsharpened pencil, one yellow plastic spoon; and one green plastic spoon. After telling the child that they were going to play a game with the toys, the tester asked the child to name the toys and their colors; if the child were unable to name a particular object or color, the tester helped him, and the child repeated the name or color after the tester. After the child was thus familiarized with the toy objects on the table, the tester presented the child with another toy (the target object) and asked the child to "Find me one on the table that is different from this one." After the child had made his choice, he was asked to justify that choice: "How are they different?" or "What makes them different?" The child's choice of object and his justification were recorded.

This procedure was followed for a total of five target objects: a red plastic square, a blue plastic comb, a yellow unsharpened pencil, a green plastic spoon, and a small, red raisin box. The set of eight choice objects and the five target objects were constructed so that each target object would have specifically appropriate counterparts in the set of choice objects. Thus, when presented with the red square, the child's set of possible choices included two objects similar in form and color and six objects different in both form and color. When the target object was the blue plastic comb, the child's set of possible choices included

an object similar in form but different in color, an object similar in color but different in form, and six objects different in both form and color. For this particular target object, there was no choice object that was similar in both form and color. When the target object was the yellow pencil or green spoon, the child had the choice of objects similar in both form and color, similar in form but different in color, similar in color but different in form, and five objects different in both form and color. There were no objects in the array similar in either form or color to the raisin box.

Several minor variations in procedure emerged during the course of the study. In addition to an occasional error in technique, there were some intentional deviations. Children who appeared upset during the session were not asked for justifications for all choices. We did, however, attempt to get a justification for the blue plastic comb target (where the child was forced to select a similar but not identical object) and for at least one other target item. Also, since several subjects decided that red squares were similar to the raisin box, we tested seven subjects using a pair of scissors as an additional target.

Results

Three independent judges, only one of whom was fully aware of the purpose of the study, were asked to judge the overall pattern of choices for whether they conformed to: (1) a pattern of consistently picking similar objects (the D & W effect), (2) a pattern of consistently correct (different) choices, or (3) a pattern which conformed to neither 1 or 2. The reason for using the pattern of outcomes was that in most instances more than one particular choice would fulfill a pattern. D & W subjects should always pick

a red square when the target was a red square, but to the blue comb target either the white comb or the blue ring would be appropriate. Children choosing correctly should never pick the red square for the red square target, but could pick the white comb, blue ring or any other object for the blue comb target.

Twenty-eight of the 29 subjects were classified without disagreement. Twenty were called D & W subjects and eight were correct. The one case involving disagreement resulted from an error of procedure in which an item was repeated. The child chose correctly on the first trial, but chose a similar object on the repetition. Since all of this child's other choices were correct, the one error was discounted. This subject was classified as correct, raising the total of that class to nine. With this correction the 29 subjects fell into a dichotomous classification in which all responses of all subjects were consistent with one of two principles.

In addition to the highly reliable judgments of the subjects' records, four of the youngest correct subjects and five of the oldest D & W subjects were re-tested one week after the initial session. Of the four correct subjects, two were again completely correct, one child made one similar choice, and one made two similar choices, but immediately corrected herself. The five D & W subjects repeated the pattern obtained earlier.

An indication that the consistency of the dichotomous classification resulted from the child's consistent use of a principle to generate his responses can be seen in the responses to two key items. To the blue comb target, 18 D & W children selected the white comb and the other two selected the blue ring. Thus, all D & W subjects selected one of the two most similar

objects when no identical object was available. The second item of interest is the response of the D & W subjects to the raisin box for which there was no intentionally similar object. Five D & W subjects selected a red square which did look similar to the raisin box and eight did not make a choice. Either of these responses is consistent with an effort to maximize similarities. Three D & W subjects were not questioned with the raisin box, so out of the 17 tested, only four subjects selected objects that were different from the raisin box. In contrast, the choices of the correct children in response to the raisin box were varied with no object being selected by more than two of the nine children.

As in Experiment I, the occurrence of the D & W pattern was imperfectly related to age. A split of the sample at the median of three years, four months, indicated that there were eight D & W subjects above the median and two correct subjects below. The probability of such a difference is .058 by a Fischer exact test. All 10 subjects three years, two months, and younger showed the D & W effect.

The choices of the correct children on the items for which there were similar objects are relevant to the formulation suggested by the data from Experiment I. The four children in Experiment I who denied the existence of a different object when forced to choose between items completely different or completely similar to the target suggested that for them "different" might require a dimension of similarity. Children in Experiment II could indicate such a belief either by a denial of an appropriate choice on certain items (e.g., the red square) or by consistently choosing a different object with a dimension of similarity when such a choice was available. Only one correct child refused to select when confronted with the raisin box, and none of the

correct subjects ever denied the existence of an appropriate choice to any items. There was, however, a pattern of choosing similar objects when they were available. For most subjects there were three targets for which a correct choice could be either a similar object or one completely different. Of the nine correct subjects, only two consistently chose a totally different object. Four subjects always picked a similar object and three more selected a similar item on two out of the three possible choices. Thus the results for seven subjects out of nine were consistent with the interpretation of "different" suggested by the four subjects in Experiment I; but none of the subjects went so far as to deny an appropriate choice. The only group of children who refused to respond with any regularity were the D & W children who failed to respond to the raisin box. There were eight such subjects. Taken in context of their other responses, however, their refusals were denials of the presense of a similar object, not the denial of a different object.

One notable result of the present study occurred in the justifications. For the D & W subjects, the blue comb forced the choice of a similar, but not identical, object. When these children were subsequently asked to justify their choice, eight of the 20 correctly did so. All eight said something to the effect that "one is blue and one is white." From the pattern of their responses it is clear that these children selected the object they did on the basis of its similarity to the target, but their justifications correctly referred to the difference. On the other items where the choice was actually incorrect, these children either referred to similarity or refused altogether to justify their choice. The correct

subjects generally were correct in their justifications for all items and referred to the dimensions of difference.

Discussion

The results of Experiment II replicate and extend the pattern seen in Experiment I, but raise further questions about interpretation. In the present version of the D & W selection task, subjects were completely reliable in using one of two principles as a basis of selection. Independent of the kinds of relations which held between the target object and the set of possible choices, most subjects (.69) chose the maximally similar object when asked for one that was different (D & W phenomenon). The remaining subjects correctly chose an object that was actually different in some way. In both groups, the particular principle employed accounted for all the responses of each of the subjects. Such clear differentiation and consistent responding, particularly among such young children, is very surprising. The facts strongly suggest, however, that the choice behavior is generated by rules that are actually related to notions of similarity and difference and not some basis unrelated to the experimenters' intent.

The data from Experiment II lead to the consideration of two difficult problems. First we must consider the relevance of the findings to our understanding of the child's use of the term "different," and second, we must consider the general problem of the relation of thought, or at least non-linguistic behavior, to language in the present experimental context.

The fact that D & W children in Experiment II are completely consistent

in using a principle which leads them to select the maximally similar object supports the conclusion that they confuse the meaning of "different" for "same" in some sense. D & W's favored hypothesis is that the children believe "different" refers to a denial of identity. That is, that the child would paraphrase "different" with "another." At first glance, however, the present data are more consistent with the notion that "same" and "different" are used as synonyms by the young subject (an alternative that D & W also mentioned) and that "different" would be paraphrased with "same." While the present data do not conclusively eliminate the identity hypothesis, the D & W subjects' consistent choice of similar items when no identical object was available suggests that the synonym explanation is simpler. It must be noted, however, that accepting the synonym interpretation does not resolve the problem since "same" is inherently ambiguous (c.f. above).

Four subjects in Experiment I suggested still another interpretation of "different." These subjects' denials of the presence of a different object when their choices were either completely alike or completely different implied that "different" requires some basis of comparison. If this were the case, a green spoon would be "different" from a yellow spoon, but not from a white comb. It is possible that such a formulation on the child's part is a variation of D & W 's identity explanation. If the child believes the request is for a "different" blue comb, for example, he might well object if no such choice were available. Alternatively, the requirement for some similarity as a basis of comparison may be an intermediate stage in the development of the meaning of "different." This latter interpretation is supported by the fact that most (seven of nine) of the <u>correct</u> subjects

in Experiment II consistently chose a "different" object that showed a dimension of similarity with the target object. Whether the "dimension of similarity" interpretation of "different" outlined above is a developmental stage sequentially related to other interpretations or represents an alternative pathway taken by individual children would seem to require longitudinal research. It should be noted that in a replication of Experiment II with 18 Johns Hopkins students, Mr. David Kaplan found that 10 subjects consistently chose a maximally different object. One student chose similar objects (the D & W effect) and justified his choice by saying that it was a different (i.e., another) red square or whatever. The remaining seven subjects were not consistent but tended to choose objects with a dimension of similarity with the target.

The final difficult question which must be briefly entertained is the relation of the children's choice behavior on the D & W task to semantics. As was noted above, there seems to be an assumption by D & W and by Clark (1970) that choice behavior in response to instructions is a fairly direct reflection of the child's understanding of the terms used in the instructions. There are two possible explanations of the D & W effect which have little or nothing to do with the semantics of "different."

First the child's response may be generated by an information processing strategy which requires him to differentiate the similar objects from the choice array before choosing from the dissimilar subset. The interuption of such a process would lead to the selection of a similar object.

A second plausible nonlinguistic explanation of the D & W effect is that the behavior may simply reflect a habit determined by the customary experience of nursery school children. When faced with a collection of objects and an adult holding another one, the child may select a similar object with no regard to the adult's specific request, simply because that is what he is usually asked to do. The clear differentiation between D & W performance may be a direct function of listening to instructions, but not of how the instructions are interpreted.

The most striking indication that a non-semantic interpretation of the D & W task must be considered is the eight children who consistently make D & W choices, but who justify a forced different choice on the appropriate basis. Thus, when forced to choose a white comb in response to the blue comb target, these subjects said correctly that one was white and one was blue to justify their choice. It is not clear to what extent these eight children illuminate the processes used by the majority of the sample, and they may represent some intermediate stage between D & W performance and correct solutions to the choice task. The other possibility, however, is that the choice behavior and justifications reflect two relatively independent processes which become connected only at some later date. At the very least the inconsistency between the choice behavior and justifications of these eight subjects suggests that there is some sort of décalage between the child's semantics and the directive functions of language.

All of these considerations suggest that at least for the terms "same" and "different" a more complex situation holds than appears to

be the case for antonymic adjective pairs and that the problem requires longitudinal study. The complexities which appear are: 1) there are some subjects who seem to use consistent principles which are different from those used by the majority; 2) these children may represent either intermediate stages or alternative stages and the question cannot be resolved by the present data; 3) for some subjects apparent semantic structure and nonlinguistic choice behavior are not perfectly isomorphic.

An interesting methodological aspect of the present study is that the behavior of individual subjects is sufficiently consistent that they may be classified with a high degree of certainty. By using a subject-by-subject analysis it is possible to avoid the distortion introduced by averaging across subjects. Also it is potentially possible to examine the development of individual subjects for sequential changes.

Study III:

A longitudinal study of young children's interpretation of "different"

Introduction

In general our knowledge of semantic development is far less sophisticated than our knowledge of syntactic or phonetic development, and the young child's use of "same" and "different" bears at least superficial similarity to one of the few regularities found in semantic development--the differential acquisition of marked and unmarked members of antonymic adjective pairs (H. Clark, 1970; Gordon, 1972; Hamilton & Deese, 1971).

It is generally believed that the acquisition of semantic features proceeds in quantum fashion adding one component at a time with the order of acquisition constrained by the structure of the semantic space (E. Clark, 1971). But McNeill and McNeill (1968) discovered a sequence in the acquisition of various forms of negation in Japanese that would appear to be related to cognitive constraints, and Greenfield (McNeill, 1970, p. 73) points to the role of individual experience in the acquisition of "dada." Thus the constraints on semantic development may come from many sources and not be limited primarily by the feature structure of the term in adult language. As is the general case, the development of the terms "same" and "different" would appear to be constrained by at least three influences. First, there is the cognitive capacity to classify, or at least see the similarities in collections; second, there are the formal linguistic properties of terms; and third, there is the individual inter-

action of the child with members of his community (see Bloom, 1970, p. 231 ff. for a similar analysis). The terms "same" and "different" are particularly interesting, however, in that their meaning is a matter of social consensus and context. As many writers have noted (e.g., Elkind, 1969) questions of similarity are inevitably post hoc. The apparently infinite categorizing ability of the human mind will allow a subject to justify a claim that any two objects are either similar or different. Also, in English the situation is doubly complex because "same" and "different" refer ambiguously to two relational concepts--identity and similarity. In the sentence, "That's the same black dog I saw here yesterday," same refers to an identity relation, while in, "Were the two dogs we saw the same color?" it refers to similarity. Adult speakers of English can generally infer from the context which relation is intended, but it is possible--or even likely--that children are not as successful.

In D & W's original paper several possible meanings of "different" were suggested but D & W favored an interpretation that "different" to a young child meant a different one of a similar type (i.e., "different" means a denial of identity). The major alternative would appear to be "different" is used as a synonum for "same." This interpretation would appear more consistent with the lexical marking approach used by Clark (1970)--though Clark himself was reluctant to extend lexical marking theory to "same" and "different."

Study II (above) indicated that the D & W effect could be reliably demonstrated with children about three-and-a-half years under a variety of conditions and the effect was relatively independent of the terms employed. Children chose similar objects when asked for one that was "same" or "like," and subjects demonstrating the D & W phenomenon also chose similar objects to "different" and "not like." In a sample of 50 children, only one subject was found who responded differentially to "not like" and "different." Children were also tested in Study II by giving them D & W choices under a variety of conditions. Target objects and choice arrays were constructed so that each subject had to choose a "different" object when the possible choices included all combinations of identical, similar, and completely dissimilar objects. In a sample of 29 children from approximately three to four years of age, all subjects could be classified as either correct on the task or as demonstrating the D & W phenomena by consistently choosing the most similar object available. Several of the D & W children refused to select when their only choice was completely unlike the target object. These findings suggested very strongly that whatever the basis of the D & W phenomenon that young children applied the rule systematically and that the phenomena did involve a rule related to similarity.

In Study II it was concluded that the synonym interpretation was most consistent with their finding that D & W subjects attempted to select the maximally similar object regardless of the nature of the available choices. There were a number of incidental findings, however, that suggested a more complex situation. Only the youngest subjects in the various D & W procedure employed, appeared simply to select the most

similar object available and then be completely satisfied with what they had done. Several of the older D & W subjects would consistently choose the most similar objects available, but then correctly refer to a dimension of difference when asked to justify their choice. Also, many subjects, when forced to choose between objects that were completely different from the target, refused to choose altogether. Some of these subjects implied that to be "different," two objects must share some basis of comparison; that is, a dimension of similarity. Finally the majority of correct subjects in the D & W procedures always chose a different object that had some basis of similarity with the target if such a choice were available. Thus, if choosing an object "different" from a yellow spoon, the subject would most likely choose a green spoon as opposed to another yellow spoon (the D & W subject's choice) or a red square.

These behaviors taken together suggested a sequence of stages in the development of the meaning of "different" that might account for all the data from the various procedures. Four stages were suggested by the data that might be sequentially related:

1) "different" means "same" with reference to similarity
2) "different" means a different one of a similar class (D & W)
3) "different" means different with some basis of similarity
4) "different" means different with reference to both similarity and identity relations

Intuitive observations made during the studies also suggested that stage 1 might be further subdivided. The most simple behavior of all might be to ignore the instructions altogether and simply match the

experimenter's object because that is what children are most likely to do in nursery school exercises. That is, for the youngest subjects, the D & W phenomenon might simply be a habit. It should be noted that while the semantics of "different" are changing through four stages in the model above, the overt choice behavior in the D & W procedure goes through only two. Children in stages 1 and 2 would select similar objects, and in stages 3 and 4, different objects.

The present study was designed to see if the model suggested above for the development of the meaning of "different" would be supported in a longitudinal study. With the cross-sectional data available from the previous studies, it is possible that the findings that are most suggestive of a number of stages are simply errors or a lack of reliability in the procedure. Also, the imperfect connection between language and non-language behavior and the deviant behaviors exhibited by a small number of subjects raised the question of whether these behaviors constituted errors or developmental stages that were intermediate between D & W and correct stages of performance. Both questions implied that a longitudinal study was imperative if the ambiguities were to be resolved.

The cross-sectional studies that had been done suggested a rather specific time for the change between the two major stages (about three years, six months) and that successive testing would necessarily be close together if more than two stages were to be seen. If this were done with identical procedures, however, a carry-over effect from one testing to the next might confound the results. These considerations

and the fact that earlier studies had suggested that the basic D & W phenomenon was highly reliable across a number of conditions led to the decision to change test materials and other procedures slightly between each testing. This would lead to greater problems of reliability, but should add considerably to the generality of the results if positive findings occurred.

Method

Subjects: Twenty-four children enrolled in a local church-operated nursery school served as subjects. The children ranged in age from 3-0 (3 years 0 months) to 3-11 at the time of the first testing. The sample was white and middle class, and none of the children had been employed in earlier studies.

Procedure: Subjects were tested three times over the school year at intervals of approximately two months. All testing sessions included one version of the D & W procedure in which the child was presented with a set of choice objects and a series of target objects. For each target object, the child was instructed to select from the choice array something "different." After most selections, the child was asked to justify his choice by answering the question, "How are these different?" Each session also included tests of other functions that were believed to be possibly correlated with the D & W phenomenon. Testing was done individually in a small separate room and all sessions were tape recorded. In the week preceding the first testing, experimenters interacted with the children in their classrooms. The specific problems employed are described below for the three testing sessions.

Session I. In the D & W procedure the choice set consisted of one red square, green spoon, yellow spoon, white comb and blue ring. The targets consisted of a blue comb, red square, and green spoon.

Additional tasks included a Piagetian free-classification task with 12 blocks of three colors and two shapes in which children were asked to "put the things that are alike together." In this task children were also asked to name colors and shapes and the use of spontaneous object-objective constructions noted (e.g., "That's a red square.").

The third task in Session I was a classification by counting problem which asked Piaget's class inclusion question without the use of the terms "same" or "more." The technique was based on the fact that most of the subjects tested could count small collections of objects. The child was presented with a collection of six objects which constituted intersecting classes. In this case, two white cubes, two green cubes and two green triangles were used. The child was asked to "count the cubes" and as soon as he finished, was asked to count the "green ones." After some delay, the questions were repeated in reverse order. For the child lacking class inclusion, the answers would be "four" to the first question and "two" to the second. More details about the free classification and classification by counting procedures will be presented in conjunction with the presentation of the results from those procedures.

Session II. Each child was retested with two versions of the D & W problem which involved two sets of objects and two procedures.

One set of objects consisted of the same objects used in Session I and the second set consisted of geometric shapes (targets: green square, white square, green circle; choice object: green square, green circle, white triangle, red circle, blue square).

One procedure was a simple replication of the standard procedure. The second procedure was the task suggested by Clark (1970) to test the identity explanation of the choice behavior. In this procedure the targets and choice objects were placed together in one collection, a target object removed from the set, the "different" question asked, and the target replaced in the set. Clark hypothesized that if the child interpreted "different" to mean a different one of a similar class of objects, he should never choose the target object. Subjects were randomly assigned to four conditions which received the two problem sets and two procedures in counter-balanced order. During the D & W procedure of Session II, children were also questioned with novel target objects that shared no dimension of similarity with any target object (i.e., completely different objects).

Session III. In Session III an extended D & W procedure with all new objects was employed. Targets consisted of a small blue car, metal bell, red spoon, green wooden block, yellow lemon and a tan plastic horse. The choice array consisted of a small blue can, a large blue car, a black plastic horse, white plastic horse, red spoon and a yellow plastic banana. In addition to the D & W procedure, the classification by counting procedure used in Session I

was repeated with a new collection of objects (purple and green plastic fruit and green plastic fruit).

Results

The data presented below will be organized around four key questions: 1) the reliability and consistency of the D & W behavior within and between sessions; 2) the relation of children's choice behavior on the D & W procedure with their understanding of the term "different" in their justifications of their choices; 3) the relation of D & W behavior to classification; 4) the possibility of describing stages in the development of "different." It should be noted that, as in previous work with the technique, the D & W procedure generates very consistent data. When a subject is classified as D & W, it indicates that all of his responses are of the most similar object available. Correspondingly, subjects who are classified as correct make no choices of similar objects. Only the behavior patterns of the "mixed" subjects must be described in greater detail. Because of the consistency of the results, the use of statistics has been kept to a minimum. Where apparent relationships occur within the data, the significance of the contingencies have been tested.

D & W Procedure (Session I): The data from the D & W procedure in Session I constituted an essential replication of earlier work with the procedure. Of the 24 children tested, 18 were perfectly consistent with either the D & W pattern of choices or were completely correct. Twelve subjects always chose the most similar object available on all

their choices (the D & W pattern) and six chose an object with a visible dimension of difference on all their choices (correct). As in earlier procedures, the breaking point between the D & W pattern and correct choices appeared to occur at about 3-6. A split of the 18 consistent subjects at the sample median between 3-6 and 3-7 found three D & W subjects 3-7 and above, nine D & W subjects 3-6 or below, and all six correct subjects abouve 3-7. The probability of this outcome occurring by chance is less than .01 by a Fischer exact test. The oldest consistent D & W subject was 3-9.

The responses of the six subjects who were not completely consistent with either the D & W or correct patterns were far from random in their behavior. One subject who was also the youngest in the sample (3-0) selected the same object on each trial. The other five subjects, however, demonstrated only two patterns of responding. Three subjects chose the red square to the red square target, the white comb to the blue comb, and the yellow spoon to the green spoon (a green spoon was also available). The other two subjects chose the blue ring to the red square, the white comb to the blue comb, and the green spoon to the green spoon. The consistency of these patterns implied that these subjects were responding on some systematic basis which was closely related to the D & W phenomenon. For subsequent analyses these five subjects plus the one who selected the same object for each choice have been classified as "mixed."

Session II: The D & W procedure on the second testing done approximately two months later demonstrated change in the frequency

of the various categories that would be expected if the D & W phenomenon were a valid developmental sequence. Table 1 gives the frequencies of outcomes on the first testing against the second testing for the 20 subjects who were tested on both procedures. The three oldest correct subjects were not retested and one previous D & W subject was not available due to illness. On the assumption that the mixed category was an intermediate developmental stage, and not simply an error, the mixed classification was retained in Table 1. The data clearly support this assumption since there is movement from D & W to mixed and from D & W to correct and from mixed to correct, but no movement in the reverse direction.

		Outcome on Session I			
		D & W	Mixed	Correct	
Outcome on Session II	D & W	7			7
	Mixed	2	4		6
	Correct	2	2	3	7
		11	6	3	20

Table 1: Outcome on Session I against outcome on Session II for the D & W procedures.

Because of the different types of materials and the two procedures used in Session II, the exact patterns of choices of the six mixed submects will not be described. Each, however, appeared to choose identical and similar objects and not randomly selected objects. The only exception to this were two subjects who responded in an unexpected fashion to the Clark procedure (see below).

The particularly interesting aspect of Session II was an attempt to conduct the experiment proposed by Clark (1970) which consisted of returning the target object to the choice array. Children who believed that "different" meant a denial of identity (that is, a different one of a similar type) should always select some object other than the target object. By and large the Clark prediction was supported, but the results were much less consistent than those found with other D & W procedures. Of seven subjects in Session II who consistently selected similar objects on both the standard procedure and the Clark experiment, five chose another similar object and not the target for each trial. One child, the youngest of the seven, selected the target on each choice, and one subject selected the target twice and another object once. The choices of the six mixed subjects were almost equally divided between correct and D & W responses on the Clark procedure and the D & W choices were divided between choices of the target and another similar object--though the majority were of the other object. Thus, it appears that most of the subjects were consistent with D & W's identity explanation of the D & W phenomenon, but the behavior on this procedure was less consistent than on other versions of the D & W task. Also,

two mixed subjects were consistely D & W on the standard procedure, but consistently chose different objects (i.e., they were correct) on the Clark version of the problem. This finding suggested that for those subjects, returning the target object to the array changed the task in some significant way. The possible significance of these results will be discussed in more detail below.

Session III: D & W data from the third testing session gave additional evidence of orderly developmental change. Of 15 subjects tested in all three sessions, seven still showed the D & W pattern. Only one of these subjects had been anything other than D & W on the earlier session. This one subject was the youngest in the sample who had been unclassified on the earlier testing due to patterns of choice that were considered deviant. Three subjects who were mixed on the third testing had been mixed on the second. Of the four correct subjects on the third session three had been correct on the earlier session and one mixed. Again, with the exception of the single deviant subject, all movement was in the direction that would be anticipated if the three categories constitute developmental stages. The fact that by the third testing the single deviant subject behaved in a fashion comparable to other subjects suggested that this subject had started from a stage developmentally prior to the other subjects.

One surprising aspect of Session III is that seven subjects still show the D & W pattern even though most are now at least 3-6 in age. It is highly speculative, but the repeated presentation of the D & W procedure may extend the period in which the child may respond in this

fashion. Even though the objects employed were changed from session to session, children talked about the previous sessions and clearly remembered what they had done before. It is possible that some of these subjects did not change to correct patterns because they did not wish to appear inconsistent.

Justifications: Study II (above) found a small number of subjects (eight of 20) who gave consistent D & W responses, but who justified a forced different choice by referring to the dimension of difference between target and choice. These data were considered suggestive of some degree of independence between the D & W choice behavior and the child's interpretation of "different." The small number of subjects showing the phenomenon also suggested, however, that the justification could be an error or chance statement. In the present study, the justification phenomenon was readily replicated, and additional information was collected concerning the selection of the phenomenon to the children's choice behavior.

The critical item for the justification question is the blue comb target. On this item, a white comb and a blue ring were available as similar choices, but no identical object was present. Twenty-one of 24 subjects chose the white comb to the blue comb target. The question posed to the child was: "How are these different?" Responses to this question for the blue comb item were examined in terms of the age of the subject, and his classification on the D & W task (D & W, mixed, correct). Choices on this item could be justified on the basis of difference ("blue, white," "That one is blue, that one is white,"

"They're different colors") or could be incorrectly justified by reference to similarity ("both combs"; "both comb hair").

The majority of all 24 subjects justified their choice of the blue comb item correctly by reference to difference and the results were closely related to age and response pattern. Table 2 gives the justification for the blue comb item for all 24 subjects classified by age and response pattern. Codes for the justifications are as follow: F = functional similarity ("Both comb hair"); S = similarity ("both combs:); D = difference ("one's blue, one is white"); and O = omission (no intelligible answer). All subjects 3-2 and below justified the blue comb item with reference to similarity. The one mixed subject 3-2 deserves special note, however, since she did justify her choice of the yellow spoon to the green spoon target by saying, "That one's green, that one yellow." Most of the subjects 3-3 and above justified the item correctly, though those who did not are with one exception D & W subjects. It should also be noted that two of the D & W subjects who supposedly justified their choice correctly may have done so for the wrong reasons. On the other items in this procedure these subjects referred only to the color of the objects. Thus, they may have correctly justified their choice on the blue comb item without actually understanding the experimenter's question.

None of the mixed or correct subjects above 3-3 made reference to similarity in their justification of the blue comb item. Table 2 indicates, however, that one mixed subject did make an error in justification on another item, and one correct subject did not justify the

Table 2

Distribution and Justifications on Blue Comb Item

as a Function of Age and D & W Category

	Category			
Age	Deviant	D & W	Mixed	Correct
3-0	F			
3-1		S,S		
3-2		S	S[1]	
3-3		D,D,D[2],F		
3-4		D[2]		
3-5			D	
3-6		O	D	
3-7		F		D
3-8		O		D
3-9		D	D[3]	D,O
3-10			D	D
3-11				D

Codes: F = functional similarity S = similarity D = difference O = omission

1 justified the choice on blue comb item with reference to similarity, but justified the choice of the yellow spoon to the green spoon correctly with reference to their colors

2 referred only to colors in all justifications and thus correctness of justification is suspect

3 justified yellow spoon-green spoon item incorrectly by saying both are spoons

blue comb item. The one mixed subject under 3-2 justified the blue comb item incorrectly, but justified a second similar choice she made correctly. Thus two subjects who are mixed on the D & W choice procedure are also mixed in their justifications.

D & W Performance and Classifications:

The writer's original belief concerning the D & W procedure was that the effect probably resulted from non-linguistic, cognitive factors. It seemed unlikely that the child's language organization could change in a radical way without an underlying change in cognitive organization. Attempts to discover non-linguistic correlates of D & W performance, however, have been consistently unsuccessful. While it is possible that the tasks that would uncover a radical change in some cognitive function at about 3-6 have not yet been employed, the child's behavior on two separate classification tasks does not appear related to performance on the D & W task.

In Session I a version of a Piagetean free-classification task was presented. A number of geometric shapes of various colors were presented and the child asked to put "the things together that go together" or put "the things that are alike together." Additional prompts were used if the child did not spontaneously group the objects with at least two collections of similar objects. If a categorization were completed the child would be asked whether he could "put them together some other way." The major source of difficulty in dealing with the classification data was the lack of variability. The majority of subjects (16) could make

a one-way classification of the objects into two groups but requests for a reclassification usually resulted in physical rearrangements of the existing groups. In Piaget's terminology (Inhelder & Piaget, 1964) the majority of subjects reverted to "graphic collections" when reclassification was requested. Of course, a rearrangement is a perfectly reasonable interpretation of "put them together some other way" even though older subjects responded almost inevitably with a reclassification.

The present sample of subjects was too young to respond to Piaget's class inclusion question in an intelligible fashion. After some preliminary attempts at such a procedure, the writer hit upon an alternative experimental technique which appeared logically equivalent to the standard class inclusion question. Most of the subjects studied in the present and previous studies of this report could count small collections of objects. The classification by counting procedure devised for the present study required the child to count intersecting classes of objects within an array. In Session I, for example, two white cubes, two green cubes, and two green triangles served as the set of objects. Twenty-one of the 24 subjects could count at least four objects and were able to generate data on the classification by counting task. Of those, eleven could count only two green objects after counting the four cubes.

Neither free classification behavior nor classification by counting was significantly related to D & W behavior. In order to maximize the possibilities of finding such a relationship, the subjects

were classified into dichotomous categories of each of the tasks. Mixed subjects on D & W were classified within the consistent group which they most closely resembled.

The criterion used for the free classification task was whether the child made at least one legitimate classification using all the objects. On the classification by counting task, the child's answer to the second question was taken as criterial. If the child counted four objects he was classified as correct; but if he counted two, incorrect. The bases of these categorizations were selected to maximize the relationships between D & W behavior and the two classification tasks. If relationships had been positive, a more thorough and conservative analysis would have been required, but neither task related significantly to D & W performance. The probability of the contingency between free classification and D & W was .33 and for classification by counting, .20 (both by a Fischer exact test).

It should be noted that the present author does not wish to suggest that the classification tests used in the present study are in any way satisfactory tests of operational classification, as Piaget would define it. These are simply two tests which are apparently related to classification, neither of which relates to D & W performance. It should also be noted that over the entire series of studies done with versions of the D & W procedure, the writer has yet to find anyting that correlates better with D & W performance than does age.

Sequential models for the meaning of different:

The overall data of the present study and incidental data from the previous studies in this report were consistent with the four stage model for the meaning of "different" proposed in the introduction to Study III. Briefly, there was the strong suggestion that children 3-2 and younger reversed the meaning of "different" for "same," that for children roughly 3-3 to 3-6 "different" referred to a different object of a similar type, that children 3-6 or 3-7 and above believed that "different" meant different with a dimension of similarity, and finally that children slightly older arrived at a meaning roughly comparable to that of adults. Because of the various pieces of data which must be brought together to make these points, data from Studies I and II will be referred to as needed. The material to be presented will be organized around the transitions between the various stages and the data which suggest the postulation of distinct stages.

The youngest children tested in the various D & W procedures seemed content to select similar objects to requests for "different" ones and did not seem bothered by the experimenter's requests for justification. Table 2 above shows this more clearly. With one exception, children 3-3 and younger are consistently D & W and justify their choices on the blue comb item with reference to similarity. The one exception is the mixed subject who justified the blue comb item with reference to similarity--the only mixed subject to do so. In Study II, a similar pattern was found. The youngest correct sub-

ject to justify an item correctly was a 3-2 D & W subject. While there were only five subjects in Study III in this youngest age range, it should be noted that there were 10 in Study II, making a total sample of 15 in the age range. In Study I, subjects 3-3 and younger were not found to reverse the meaning of "different" in the judgment task, but with one exception all children in this age range employed one of the response generating heuristics discussed in that study. Thus, the data from Study I were not inconsistent with the first stage of D & W performance being postulated, but did not support it directly either.

At the present time there is only impressionistic data to support the notion suggested above that the first stage of D & W performance might be divided into two substages. The automatic character of the response of many of the younger subjects suggested that they might not be listening to the information at all and simply responding on the basis of a habit. A few subjects began to make their choice prior to the experimenter's instructions and such behavior supports a non-linguistic interpretation of the behavior. There is, of course, nothing in the behaviors recorded in the D & W task (choices and justifications) that compels a distinction.

The next question is what data suggest a differentiation at approximately 3-3, when the choice behavior on the D & W task did not change until approximately 3-6. The justifications are clearly the most compelling, but other features of the data are consistent with the assertion. Table 2 indicates that three of four subjects 3-3 (all of whom are D & W)

correctly justified the blue comb item. There are only two subjects 3-3 or older that make D & W choices and fail to justify the critical item correctly. Comparable data were seen in Study II where proper justifications were generally made by the older subjects.

The youngest consistently correct subject seen in the various studies with the D & W procedure was 3-3. By and large, however, the change in choice behavior occurred at about 3-6 or 3-7. In Study III, for example, the youngest correct subject is 3-7. If we examine the subject in the intermediate age range of Study III we see that only two of 10 subjects make D & W choices and justify the blue comb item choice on the basis of similarity. These two subjects are acting like younger subjects, and the majority of subjects in the 3-3 to 3-6 range are either mixed or are D & W who justify correctly.

The data from the experiment proposed by Clark (see above under Session II) is relevant at this point. If the proposed stage structure is correct, subjects under 3-2 would select the target objects and older D & W subjects would select other similar objects. Unfortunately, the data directly relevant to the first half of this prediction are generated by only a single subject, but that subject is consistent with the prediction. The one D & W subject who picked target objects on each choice of the Clark experiment was the youngest D & W subject (3-3 at the time of the testing). Five of six older D & W subjects selected similar objects on all trials, and the remaining subject selected some targets and some similar objects.

The one subject younger than 3-3 who generated data on the Clark experiment made all correct choices when the target object was returned to the array and D & W choices when the target was not returned. This suggests that this subject was also operating on some principle that differed from that held by older subjects.

The most obvious of the stage changes in the D & W data occurs at about 3-6 to 3-7. At this point children start correctly choosing objects that differ on at least one dimension from the target. Several lines of evidence suggest, however, that the youngest correct subjects still hold a theory of "different" that deviates in several ways from that held by older children.

The earliest indication that young correct subjects believed that to be different implied a dimension of similarity occurred in Study II, Experiment I above. Four subjects who refused to select objects in response to the requests for something "different" volunteered the opinion that they needed a similar object to fulfill the request. In light of their refusal to respond, it is impossible to classify these four subjects into D & W or correct, but their statements implied that they would have been correct if they had made choices.

The second line of evidence is the correct subjects in Study II, Experiment II. None of the correct subjects in this condition denied the existence of an appropriate choice on any item, but they did tend to choose similar objects when such choices were available. Out of nine correct subjects, seven chose a different but similar object on

most trials where such a choice was available. The remaining two subjects chose completely different objects on these trials.

The data from correct subjects in Study III are also consistent with the theory that "different" implies for a time a dimension of similarity. By using longitudinal sampling, it is possible to compare the number of different but similar objects chosen by subjects on two different testings. The proposed theory of the meaning of "different" makes the straight forward prediction that subjects will select fewer similar objects on the second testing. Because of changes in technique, subjects who were not retested, etc., there are only a few subjects who generate useable data. Those that are available, however, are completely consistent with the predicted changes. There are six subjects who were correct on one session and who were retested in the next session. All six were still correct at the second testing. Of the six, four showed a decrease in the number of similar, but different, objects selected on the trials where such a choice was available and the other two did not pick similar objects on either testing. Thus, all the available data are consistent with the notion that when they first correctly begin selecting different objects, the children try to find choices that have some basis of similarity with the target. The data from Studies I - III are admittedly from a small number of subjects, but it must be noted that they are almost completely consistent with the prediction. In Study III there are no exceptions out of six subjects who generate useable data.

What is not revealed by the present method of study is what meaning the oldest subjects attribute to the term "different." The paragraphs above strongly suggest that when children first start choosing correctly in the D & W task, they still hold a theory of "different" that is not equivalent to that held by older children. There are no available data to be derived from the D & W task, however, that tell us how those children differ from still older children or adults.

Conclusions

The studies reported above have been presented in highly detailed fashion with emphasis on the descriptions of individual subjects. It is difficult and time-consuming to follow such accounts, and yet, this style of reporting is essential to the major points to be made. Developmental changes are rarely if ever perfectly correlated with age and simply comparing the average outcomes of two groups of differing ages is a suspect procedure which leads to almost certain distortions of the processes studied. Averaging also eliminates the possibilities that different developmental pathways to the same end result can be discovered. In semantic studies, for example, current linguistic theory (McNeill, 1968; E. Clark, 1971) suggests an invariant pattern of development constrained by formal aspects of linguistics. Individual children, however, may deviate from the expected pattern in consistent and stable ways. Averaging across groups would eliminate the possibility of detecting such subjects. The work of Bloom (1970) strongly suggests that there can be systematic individual differences in something as basic as the strategy of early grammar acquisition.

Although they were not originally conceived in this fashion, the present studies became developmental studies of the meaning of "different." The original intent of the investigator was to study the interactions between cognitive constrains and social inputs as they were reflected in language. The data, however, forced a reevaluation of this plan. The outcomes of the series of studies reported above seem to lead to a fairly simple conclusion: The deviant choice behavior seen in response

to instructions involving the term "different" appears to be a fairly direct reflection of the semantics of the term, and the semantic structure changes in an orderly fashion that is developmentally stable and closely related to age. Moreover, the process can be described as a series of four discrete stages which are logically related to the semantics of the term.

In Stage I, the child seems to recognize that "different" refers to relations of either similarity or discrepency, but he apparently believes that it refers to the positive aspect of the relations. That is, "different" means "same." It is extremely unlikely that the child knows the different relations to which the term may refer or can discriminate the proper referents from the context.

In Stage II, the child makes his first attempt at being systematic with the use of the term. In the D & W choice task, the child is still 100% wrong in his choices, but several lines of evidence suggest that he holds an intellectually defensible theory. "Different" in terms of this theory means a different one of a similar type of object. That is, "different" is a denial of an identity relation. The child in this stage consistently picks the most similar object available whether it is completely similar or differs on a dimension. Thus, he appears to be operating on class membership defined by one or more dimensions and selects the maximum overlap of dimensions. As long as there is one dimension of similarity, it is possible for the child to select a "different" one of that class. Children in this stage protest the request for a "different" object, however, if there is no similar object available. Across the

several studies performed, children who would normally select similar objects would refuse to make any selection if their choices were completely unlike their target. Thus, any dimension of similarity was sufficient to define a class such that a different member of the class could be selected, but with no dimension of similarity the task was undefined. It is, therefore, not surprising that when the child's choice behavior changes to being correct, he still prefers objects with a dimension of similarity.

In Stage III, children consistently chose objects with a dimension of difference, but also chose those which shared some basis of similarity. The consistency of this finding suggests that Stage III children have not yet differentiated their use of the term "different" from the need for a physical referent. The question of "different" for these children must be defined with regard to some visible characteristic. It is only later in Stage IV that the children come to realize that objects that are completely different are, in fact, different.

Unfortunately there is nothing in the present series of studies to determine the impetus for the observed changes in the meaning of "different." The changes are orderly and closely related to age, but the problem of why the child redefines his terms in this manner is open to question. The process by which this occurs in probably close to that discussed by Vygotsky (1960) in his discussion of the acquisition of concepts. Children know something of the adult meaning of concepts by virtue of the referents of the terms. Because of the fact of common

referents, children are able to converse with adults and reveal the discrepancies that exist between themselves and adults only under exceptional circumstances. One of these exceptional circumstances is the D & W task. When the experimenter asks for similar objects, there is no evidence that anything is amiss. When the question is "different," however, the young child is completely wrong.

Although no correlation was found between D & W performance and classification, one still might presume that the changes observed result from changes in non-linguistic cognitive functions. The most likely cognitive change that might produce the observed data is the necessity for external referents. For adults, most relational terms lack physical referents, but children may be constrained in their use of terms to those instances where there are such referents. In fact, for the youngest children "same" and "different" may not be relational terms. Often young children say "that's the same and that's the same." As D & W suggest, this may be a shortened form of "that's the same color," where "same" is used as a label rather than a relation. Stages II and III described above would support the belief that the events observed in the development of the use of "different" represent a gradual divorcement from physical referents and an increase in abstract relational thinking.

References

Bloom, L. Language development: Form and function in emerging grammars. Cambridge, Mass.: The M.I.T. Press, 1970.

Clark, E. V. On the acquisition of the meaning of Before and After. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 266-275.

Clark, H. H. The primitive nature of children's relational concepts. In J. R. Hayes (Ed.), Cognition and the development of language. New York: Wiley, 1970.

Cole, M., & Bruner, J. S. Cultural differences and inferences about psychological processes. American Psychologist, 1971, 26(10), 867-876.

Donaldson, M., & Balfour, G. Less is more: A study of language comprehension in children. British Journal of Psychology, 1968, 59, 461-472.

Donaldson, M., & Wales, R. J. On the acquisition of some relational terms. In J. R. Hayes (Ed.), Cognition and the development of language. New York: Wiley, 1970.

Elkind, D. Conservation and concept formation. In Elkind & Flavell (Eds.), Studies in cognitive development. New York: Oxford University Press, 1969.

Gordon, J. Conservation and the marked semantic distinction: Indications of a thought-language interaction. Unpublished doctoral dissertation, The Johns Hopkins University, 1972.

Greenberg, J. Language universals. The Hague: Mouton, 1966.

Greenfield, P. M., & Bruner, J. S. Culture and cognitive growth (revised version). In D. A. Goslin (Ed.), Handbook of socialization theory and research. Chicago: Rand McNally, 1969.

Griffiths, J. A., Shantz, C. A., & Sigel, D. E. A methodological problem in conservation studies: The use of relational terms. Child Development, 1967, 38, 841-849.

Hamilton, H. W., & Deese, J. Does linguistic marking have a psychological correlate? Journal of Verbal Learning and Verbal Behavior, 1971, 10, 707-714.

Hays, W. L. Statistics for psychologists. New York: Holt, Rinehart, and Winston, 1963.

Inhelder, B., & Piaget, J. The early growth of logic in the child. New York: W. W. Norton & Co., 1964.

Kofsky, E., & Osler, S. F. Free classification in children. Child Development, 1967, 38, 927-939.

McNeill, D., & McNeill, N. B. What does a child mean when he says "no"? In E. M. Zale (Ed.), Language and language behavior. New York: Appleton-Century-Crofts, 1968.

McNeill, D. The acquisition of language: The study of developmental psycholinguistics. New York: Harper and Row, 1970.

Piaget, J. The moral judgment of the child. New York: Free Press, 1965.

Vygotsky, L. S. Thought and language. Cambridge, Mass.: The M.I.T. Press, 1960.